

*Citation for published version:*

Denham, EL 2020, 'The Sponge RNAs of bacteria - How to find them and their role in regulating the post-transcriptional network', *BBA - Gene Regulatory Mechanisms*, vol. 1863, no. 8, 194565.  
<https://doi.org/10.1016/j.bbagr.2020.194565>

*DOI:*

[10.1016/j.bbagr.2020.194565](https://doi.org/10.1016/j.bbagr.2020.194565)

*Publication date:*

2020

*Document Version*

Peer reviewed version

[Link to publication](#)

*Publisher Rights*

CC BY-NC-ND

**University of Bath**

**Alternative formats**

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# **The Sponge RNAs of bacteria – how to find them and their role in regulating the post-transcriptional network**

Emma L. Denham

University of Bath, Department of Biology and Biochemistry, Claverton Down, Bath, BA2 7AY

[e.l.denham@bath.ac.uk](mailto:e.l.denham@bath.ac.uk)

@gingermicrobe

+44 1225 383424

## **Keywords**

RNA sponge, sRNA, degradation, post-transcriptional regulation, network, bacteria

## **Abstract**

In bacteria small regulatory RNAs (sRNAs) interact with their mRNA targets through non-consecutive base-pairing. The loose base-pairing specificity allows sRNAs to regulate large numbers of genes, either affecting the stability and/or the translation of mRNAs.

Mechanisms enabling post-transcriptional regulation of the sRNAs themselves have also been described involving so-called sponge RNAs. Sponge RNAs modulate free sRNA levels in the cell through RNA-RNA interactions that sequester (“soak up”) the sRNA and/or promote degradation of the target sRNA or the sponge RNA-sRNA complex. The development of complex RNA sequencing strategies for the detection of RNA-RNA interactions has enabled identification of several sponge RNAs, as well as previously known regulatory RNAs able to act as both regulators and sponges. This review highlights techniques that have enabled the identification of these sponge RNAs, the origins of sponge RNAs and the mechanisms by which they function in the post-transcriptional network.

## Introduction

### sRNAs - a flexible regulator

The first demonstration of a small RNA that regulates an mRNA transcript by RNA-RNA base pairing in a prokaryote was MicF in *Escherichia coli*. Mizuno et al [1] showed it to regulate the mRNA for the outer membrane porin OmpF through annealing of the two RNAs. Since this discovery, our understanding of the richness of the world of sRNA activity in post-transcriptional gene regulation has been accumulating. Importantly MicF has both multiple mechanisms of action and multiple targets [2]. For example, the 5' end of MicF blocks formation of the initiation complex during *lrp* translation and therefore translation of the transcription factor Lrp is reduced [3]. The same region of MicF also pairs with a sequence near the 5' end of the *yahO* transcript which likely promotes exonuclease degradation of this mRNA [4]. Alternatively, by binding to the coding sequence of *lpxR*, MicF targets the mRNA for endonuclease cleavage [4].

Our understanding of how many different mechanisms exist by which sRNAs can tune gene expression has expanded with each sRNA that has been studied. Other examples of regulatory mechanisms include prevention of RNase degradation as was shown for the sRNA RoxS of *Bacillus subtilis* [5]. RoxS binds to the extreme 5' end of the *yflS* transcript blocking exoribonuclease activity. Additionally, RoxS is present in two forms, the full length and a processed form, that have different target specificity [6]. Although some sRNAs have been described that control the expression of a single mRNA target, such as RnaC of *B. subtilis* (which controls the heterogeneous translation of the global transcription factor AbrB [7]), many sRNAs have been described that control large regulons of mRNAs. One of the best described is GcvB, an sRNA conserved in the Enterobacteriaceae that is involved in global control of amino acid metabolism, controlling the level of 45 different mRNAs [8]. The converse is also true, single mRNAs can be controlled by multiple sRNAs depending on the environmental condition. The sigma factor  $\sigma^S$  encoded by the *rpoS* gene controls the general stress response in *E. coli*, the regulon of which includes over 400 genes, including 4 sRNAs SdsR, SdsN, GadY, and SraL [9]. The translation of the *rpoS* transcript is positively controlled by three different sRNAs under different conditions DsrA, ArcZ, and RprA [9].

## **How to control unrequired sRNAs – identification of sponge RNAs**

As with every other gene, it is to be expected that the cell will require regulatory mechanisms to also control the levels of sRNAs. However, historically, sRNAs have been analysed from the perspective of which mRNAs they are able to bind to. Bioinformatic target prediction tools such as CopraRNA [10, 11] and TargetRNA [12, 13], are restricted to mRNA binding sites and are generally constrained to the 5' region of mRNAs. Over the past few years, however, a growing number of examples have been described where sRNAs interact with each other, changing the levels of targets normally regulated by one or both of the sRNAs.

This sRNA-sRNA interaction forms the basis of the competing endogenous RNA (ceRNA) hypothesis that was first described for eukaryotic RNAs [14]. MicroRNAs (miRNAs) are short 21-23 nucleotide RNAs expressed by eukaryotes that can recognise hundreds of mRNA targets. The term “sponge RNA” was coined by Ebert et al in 2007 who showed that expression of an engineered RNA, that could base pair with the seed region of an miRNA, could compete with the targets of miRNAs for binding [15]. This in turn removes the miRNA from the regulatory pool by titration, so alleviating the regulation of the miRNAs native target.

The first natural sponge RNA was identified at the same time in *Arabidopsis thaliana* [16]. The non-protein coding gene IPS1 was shown to contain a motif with a sequence that was complementary to the phosphate starvation induced miRNA miR399. The interaction between the two RNAs results in miR-399 being sequestered, which resulted in increased accumulation of the miR-399 target PHO2 mRNA and subsequently reduced phosphate content in the shoot [16]. Franco-Zorrilla et al used the term “target mimicry” to define this mechanism of post-transcriptional regulation [16]. Since then many different examples of sponge RNAs have been discovered in eukaryotic species [17]. Sponge RNA function has not only been shown for miRNAs, but also for long non-coding RNAs, pseudogene transcripts, circular RNAs, viral RNAs as well as protein-coding transcripts [18]. It has also been shown that some sRNAs are able to directly titrate the level of certain proteins, thereby modifying the proteins' activity. This is an analogous function to RNA-RNA sponges (for review see [19-21]).

The first bacterial RNA shown to fit the ceRNA hypothesis and act as a sponge RNA was described in *Salmonella enterica*. Both *Salmonella* and *E. coli* are able to use the breakdown products of chitin, chitobiose and chitotriose. The gene encoding the membrane protein porin ChiP, that allows these chitosugars to diffuse across the outer membrane into the cell, is transcribed at a basal level, but is post-transcriptionally repressed by the constitutively expressed sRNA ChiX [22]. When chitosugars are present, the *chb* operon is transcriptionally upregulated and the suppression of ChiP is relieved by the binding of ChiX to the intercistronic spacer sequence between *chbB* and *chbC* [22] (Table 1).

Several reviews [19, 23-25] have concentrated on the three best characterised sponge RNAs: the intergenic region of the *chbBC* transcript [22, 26], SroC [27] and the external transcribed spacer of the Leu-tRNA, called 3'ETS<sup>leuZ</sup> [28]. Many other RNA molecules have, however, been described that can regulate sRNAs (summarised in Table 1). Three points arise from the comparison of these RNAs. First, the RNAs have been identified using a number of different experimental techniques, second, they arise from different origins and third, they function by a wide variety of mechanisms. This has resulted in many terms being used to describe sponge RNAs. This review will focus on these three topics.

## Techniques to identify sponge RNAs

In the early 2000s, pioneering screens in *E. coli* led to the identification of many sRNAs through the use of both microarrays and computational predictions [29-31]. This accumulated evidence provided the first indication that bacterial sRNA regulatory networks might be more complex than the transcription factor, sigma factor and signalling systems that had been identified to date. The introduction of tiling arrays led to the identification of numerous RNA species that were not in the original genome annotations. For example, the use of tiling arrays to map the transcriptional activity of *B. subtilis* in 104 different biological conditions, identified over 1500 previously unannotated condition-specific regions [32]. Many of these regions encode independent transcripts without protein coding function. These included antisense RNAs and previously unidentified untranslated regions (UTRs) at both the 5' and 3' ends of transcripts. This technique also enabled the identification of RNAs

important for pathogenesis in *Listeria monocytogenes* [33] and *Staphylococcus aureus* [34]. The introduction of the now routine RNA sequencing techniques RNA-seq [35], dRNA-seq [36], Term-seq [37] and Rend-seq [38] has allowed whole transcriptomes to be mapped to single nucleotide resolution and can distinguish primary and processed transcripts. Close analysis of these RNA-seq data sets from the Enterobacteriaceae led to the identification of the sponge RNA SroC [27] (Table 1). This RNA had been identified in multiple transcriptome experiments as a stable fragment from the *gltIJKL* operon mRNA. Despite being operonic, Miyakoshi et al [27] showed this to be produced through an RNase E processing event from a transcript containing only the *gltI* coding sequences. In turn they showed that this fragment is a negative regulator of the sRNA GcvB [27], which is itself also a negative regulator of *gltIJKL*, preventing its translation [8, 39].

The RNA-seq methods that have been used to identify sponge RNAs are summarised in Table 2 and the sponge RNAs that have been identified are described below.

### **Protein based methods for the identification of Sponge RNAs**

Many techniques have been developed to enable the identification of RNAs associated directly with proteins. Hfq is a conserved RNA binding protein that attaches to many sRNAs in the *Enterobacteriaceae*, helping sRNAs to find their mRNA targets [40]. Studies with Hfq have proved a powerful means to identify not only sRNAs and their biological roles, but also the mRNA targets of sRNAs.

Co-immunoprecipitation of RNA binding proteins such as Hfq [41] is the simplest technique that has been used to identify RNAs associated with the protein (Table 2). Indeed, it was the combination of SroC being identified as a stable RNA fragment of the *gltIJKL* mRNA and its interaction with Hfq that led Miyakoshi et al to hypothesise that this RNA may have a regulatory function [27]. Since the first investigation in *S. enterica*, others have identified hundreds of Hfq associated sRNAs and their RNA partners.

A methodological improvement to co-immunoprecipitation was brought about through the introduction of a crosslinking step to stabilise the RBP-RNA interactions (Table 2). The cross-linking step enables not only the identification of the RNAs bound to the specific protein

being studied, but also the mapping of the specific interaction sequences within individual RNAs. CLIP-seq (Cross-linking Immunoprecipitation) was the first technique to improve on the basic method of co-immunoprecipitation of RNAs bound to an RBP and many variations of this method have been developed [42]. CRAC (Crosslinking and Analysis of cDNA) is one variation (Table 2) and was carried out on Hfq in an EHEC *E. coli* 0157:H7 strain [43]. During CRAC a bipartite tag is added to the RNA-binding protein of interest and following UV cross-linking, the RNA-protein complex is purified under highly denaturing conditions. Hfq CRAC identified several sRNAs that are transcribed from prophages, AgvB and AsxR. These two sRNAs were shown with further experiments to act as sponge RNAs for sRNAs encoded by the host chromosome; AgvB interacts with GcvB, while AsxR interacts with FnrS (Table 1).

Although co-immunoprecipitation is a powerful technique for identifying RNAs that bind protein, it is limited to providing a list of bound RNAs, but no information as to which two RNAs are themselves interacting. A powerful advancement in RNA-protein immunoprecipitation techniques was the development of methods such as RIL-seq (RNA interaction by ligation and sequencing) [44] and CLASH (Cross-Linking, Ligation and Sequencing of Hybrids) [45] (Table 2). These techniques not only identify which RNAs are bound to the RNA binding protein, but also that enable mapping of RNAs that are interacting with each other. Several studies have exploited techniques where chimeras are formed between the interacting RNAs that are captured *via* their interaction with the Hfq protein [44, 45]. Proximity ligation uses an RNA ligase to join the ends of two RNAs that are interacting. In RIL-seq and CLASH this takes place after digestion of the RNA that is not protected by being bound by the protein. This in turn allows the identification of the particular RNA sequences involved in the interactions.

RIL-seq identified PspH, a sponge RNA present in the 3' UTR of *pspG* that reduces the availability of the sRNA Spot 42 (Spf) for interaction with its targets (Table 1) [44]. Remarkably, a large-scale CLASH experiment carried out over the entire growth curve of *E. coli* identified over 100 sRNA-sRNA interactions [45]. Forty sRNA-sRNA interactions were also mapped in an RNase E CLASH data set [46], only a quarter of which overlapped with those identified in the Hfq CLASH dataset. The ArcZ-CyaR interaction was studied in depth during the Hfq CLASH study [45] (Table 1). Moreover, an interaction between CyaR and



GcvB was also identified substantiating even further the existence of a network of sRNA-sRNA regulatory interactions. These experiments hint of the richness of the RNA-RNA network and suggest that we have only just scratched the surface of understanding how post-transcriptional regulation is implemented in the cell.

### **Non-protein-based methods for the identification of sponge RNAs**

The power of using an RNA binding protein as a scaffold to identify RNA interactions is undeniable. As described above studies of Hfq in the Enterobacteriaceae have enhanced not only mapping of the post-transcriptional regulatory network, but also our understanding of how it functions. However, in Gram-positive bacteria such as *B. subtilis* and *S. aureus* no RNA binding protein has been identified to play the equivalent role of Hfq within the Enterobacteriaceae, even though Hfq is conserved in these species. Deletion of Hfq in *B. subtilis* or *S. aureus* has little or no effect, and no role in sRNA interactions have been established [47-50]. This suggests that many RNA-RNA interactions are either protein independent or the protein mediating their interaction has yet to be identified. Therefore, alternative strategies are required to identify RNA-RNA interactions that are either protein independent or where the protein is unknown.

A modified CLASH protocol has been developed to enable the identification of interacting RNAs without the requirement for a protein bait (Table 2). This method is advantageous for bacteria where genetic manipulation is not feasible as it does not require the use of RNA tagging or protein expression. Using the psoralen 4'-aminomethyl trioxsalen (AMT) and UV irradiation at 254 nm it is possible to create reversible crosslinks between interacting RNAs. Once the RNA has been crosslinked, extracted and single stranded portions of the RNA removed, the RNA is ligated to form chimeras between the interacting RNAs. The crosslinking is reversed by exposure to UV irradiation at 365 nm and chimeric RNAs are identified by RNA-seq.

Three versions of this technique have been developed and used in eukaryotes [51-53]. In prokaryotes the technique has been implemented in the model systems *E. coli* [54] and *B. subtilis* [55]. Although many novel RNA interactions were identified in *E. coli*, none involved sRNAs. However, the study in *B. subtilis* found both known and novel interacting partners

for the two sRNAs RoxS [5, 6] and FsrA [56, 57], and a sponge RNA called RosA that interacts with both RoxS and FsrA (Table 1).

There are several limitations of this method that could be solved in the future. As AMT only forms inter-strand crosslinks between uridine residues, if uridines are not present in close proximity to the interacting RNA stretches, the interaction is unlikely to be captured. The efficiency of the proximity ligation step of RNAs that are captured by the crosslinking is also low and results in ligation of both interacting and non-interacting RNAs, causing noise in the data set. However, as shown for *B. subtilis*, it does identify validated RNA-RNA interactions and therefore represents a further technique to map RNA-RNA interactions [55].

GRIL-seq (Global small non-coding RNA target identification by ligation and sequencing) is an alternative method that produces chimeric RNAs *in vivo* through the expression of an inducible RNA ligase [58, 59] (Table 2). GRIL-seq can be carried out on both individual sRNAs and also on the global RNA pool using HI-GRIL-seq (high-throughput). GRIL-seq was developed in *Pseudomonas aeruginosa* where the iron regulated sRNA PrrF1 was overexpressed concomitantly with the RNA ligase [58]. During this investigation, the 3' end of the *kata* gene was identified as interfering with the negative regulation of *kata* by PrrF1, suggesting that *kata* 3' is a PrrF1 sponge (Table 1). Hi-GRIL-seq has the advantage that it requires no prior knowledge of sRNAs within the cell, but still involves the expression of the RNA ligase (Table 2). Novel RNA-RNA interactions were identified in *P. aeruginosa* by Zhang et al, although no specific sponge RNAs were reported [59].

While potentially revealing, GRIL-seq and Hi-GRIL-seq are likely to be limited in the kinds of interaction they can identify. The modified CLASH protocol and GRIL-seq both use proximity ligation. However, during the modified CLASH protocol ligation takes place *in vitro*, whereas for GRIL-seq this is *in vivo*. The modified CLASH protocol removes the single stranded RNAs not involved in the interaction, therefore increasing the likelihood that the ligation is made between two RNAs that are physically interacting, whereas the removal of non-interacting or single stranded RNA cannot take place *in vivo*. Therefore, a potential technical issue with GRIL-seq compared with the modified CLASH protocol, is that this lack of a step to remove the single stranded parts of the RNA is likely to have an influence on the free ends of the

two interacting RNAs finding each other. This in turn is expected to limit the kinds of interactions that can be identified.

Interactions involving individual sRNAs can be identified through the use of RNA tags. The MS2 RNA tag has been used regularly to identify RNA-RNA and RNA-protein interactions [60] (Table 2). The use by the Massé group of MS2 affinity purification coupled with RNA sequencing (MAPS) has allowed the mapping of the targetomes of several sRNAs in several species of bacteria including *E. coli* [28, 61] and *S. aureus* [62, 63]. This technique involves adding the MS2 RNA tag sequence to an RNA of interest. The interaction between the MS2 tag and the MS2 RNA binding protein is very strong and therefore interacting RNAs can be greatly enriched through the purification process. MAPS was applied to the *E. coli* sRNAs RyhB and RybB, and not only identified new mRNA targets, but also resolved the 3' external transcribed spacer (ETS) of the *glyW-cysT-leuZ* polycistronic tRNA precursor as a *bona fide* interaction partner [28]. The expression of the *glyW-cysT-leuZ*, and therefore concomitantly the 3'ETS<sup>leuZ</sup>, is constitutive. This sets an expression threshold that RyhB and RybB must reach before there is a sufficient concentration of these sRNAs for them to affect the levels of their specific targets (Table 1). Until this expression threshold has been met, the 3'ETS<sup>leuZ</sup> acts as a sponge RNA to prevent RyhB and RybB interacting with their respective mRNA targets [28].

With the increased use in sequencing as a routine experimental practice, and the improvement of methods and sensitivity to identify subsets of RNAs, the identification of interacting RNAs including RNA sponges is only likely to increase.

### **Mechanisms that produce sponge RNAs**

As has already been highlighted by the description of *gltIJKL*, GcvB and SroC, and ChiX, *chiP* and the intergenic spacer region of the *chb* operon, some bacterial sponge RNAs are part of functional RNAs that are downregulated by an sRNA under certain conditions, but that switch to functioning as a sponge RNA following their accumulation, often in response to environmental cues [19]. As highlighted above, the use of large sequencing-based studies that capture RNA-RNA interactions has expanded not only the network of interactions, but

also the repertoire of RNA sequences that can act as regulatory RNAs. Table 1 indicates there are two groups of bacterial sponge RNAs; processed mRNAs and independent transcripts (those that are controlled by their own promoter and terminator sequences). The majority of sponge RNAs have been identified in the Enterobacteriaceae. This is primarily due to the identification of the RNA binding protein Hfq, but also because of the ease of genetic manipulability offered with these organisms. This has led to a biased view as to the mechanisms by which sRNAs function. Figure 1 depicts the mechanisms by which sponge RNAs have been shown to be produced, and several possible mechanisms by which they could be produced, but have yet to be demonstrated in the literature.

### **Independently transcribed sponge RNAs**

A gene is transcribed through the recognition of a promoter sequence by RNA polymerase core enzyme bound to a sigma factor and, in some cases, activated by specific transcription factors. Transcription continues until a terminator sequence is encountered. This is how the majority of sRNAs that have been identified to date are expressed. That is, they are primary transcription products. This is also true for a group of sponge RNAs identified in both Gram-positive and negative species (Figure 1a). In *B. subtilis* the independently transcribed sRNA RosA is regulated by the carbon catabolite control protein CcpA and has the capacity to interact with other functional sRNAs [55] (Table 1). Four independently transcribed sRNA sponges have been described in *E. coli*: AgvB and AsxR, from the prophages of EHEC 0157:H7 [43], and ArcZ and CyaR from the host chromosome [45]. Intriguingly, ArcZ [45, 64] and CyaR [65] have been shown to have mRNA targets, in addition to their direct role in sRNA regulation by sponge mechanisms. This suggests a division between sRNAs that only act as sponge RNAs and those that regulate both mRNA and sRNA targets.

### **Processing transcripts to release regulatory fragments**

Many sRNAs have been shown to come from the processing of transcripts (Figure 1b, d, e). These processed transcripts, that are not immediately turned over by degradative RNases, have added extra complexity to our understanding as to how the sRNA landscape is produced and controlled. RNAs with regulatory activity have been shown to be generated from different areas of transcripts, including intergenic regions, intergenic spacers of tRNAs and 3'UTRs.

### Intergenic spacer regions

The transcription of different isoforms of the same gene or operon can enable different functions of an RNA. One example of this is the *gltIJKL* operon of *Salmonella* that is transcribed in two forms: a full-length mRNA that encodes for the glutamate/aspartate ABC transporter and a transcript that ends at a Rho-independent terminator situated between the *gltI* and *gltJ* genes (Figure 1b). The sRNA SroC is produced through an RNase E processing event of this transcript to release the intergenic spacer region between *gltI* and *gltJ* [27]. This RNA was initially thought to be an RNA degradation product [27, 39], however SroC is an Hfq associated sRNA [41, 43, 66] with a well described regulatory function [27]. The translation of the *gltIJKL* operon is prevented by the sRNA GcvB. However, SroC is a sponge RNA for GcvB and the interaction between the two sRNAs results in derepression of the GcvB regulon [27].

As previously mentioned, in the Enterobacteriaceae the ChiX sRNA negatively regulates expression of the gene encoding the chitosugar outer membrane porin *chiP* that is expressed at a moderately high basal level. Binding of ChiX causes the degradation of the *chiP* mRNA target, while the sRNA is recycled rather than being degraded. The chitosugar inner membrane transporter and enzymes required for chitosugar metabolism encoded by the *chb* operon are regulated by the presence of the chitosugar through alleviation of the transcription repression by NagC. The intergenic spacer region between *chbB* and *chbC* acts as an sRNA decoy through its ability to form a complex with ChiX (Figure 1c). The decoy sequence is part of the entire *chb* transcript and it appears that, at the point of binding to ChiX the entire transcript is still present. However, immediately after the two RNAs interact, RNase E cleaves the *chb* operon mRNA into a 400 nt intermediate [22, 26].

### Intergenic spacers of tRNAs

The sponge RNA 3'ETS<sup>*leuZ*</sup> is produced from RNA processing of the *glyW-cysT-leuZ* polycistronic tRNA transcript into the mature forms of these three tRNAs. The ETS spacer downstream of the *leuZ* tRNA is removed and forms a stable sRNA (Figure 1d) capable of interacting with both RyhB and RybB [28] (Table 1). At present it is unclear whether this is a conserved mechanism found in other tRNA intergenic spacers or other species. The study by

Lalouna *et al* also showed that the *metZ-metW-metV* tRNA operon transcript was able to interact with both RybB and MicF sRNAs. However, in contrast to the 3'ETS<sup>leuZ</sup>, these intergenic spacers are not matured to stable sRNA fragments, suggesting that if these sequences also act as sponge RNAs, they use alternative mechanisms [28]. The authors also showed that for the Enterobacteriaceae, some 3' ETS sequences are similarly or even more conserved than the 3'ETS<sup>leuZ</sup>. In particular, the 3'ETS<sup>asnW</sup>, 3'ETS<sup>argQ</sup>, 3'ETS<sup>metY</sup>, and 3'ETS<sup>valW</sup> all showed great sequence conservation and the length of the regions is sufficient for potential base pairing [28].

### 3' UTRs

The SkatA sRNA is produced by processing the 3' UTR sequence from its parent mRNA, and is similar to the processing of the SroC sRNA from the *gltIJKL* transcript [58] (Figure 1e). Mapping of the *P. aeruginosa* transcriptome had previously only identified one promoter for the *kata* transcript from which the SkatA RNA originates. Therefore, this sRNA is most likely generated from an endonucleolytic cleavage event of the native *kata* mRNA, although the RNase involved has not been identified. Another sponge RNA that is derived from a 3' UTR is PspH from the *pspG* gene. However, PspH most likely results from a promoter present in this region rather than endonucleolytic cleavage [44] (Figure 1f). Analysis of transcriptomes has led to the identification of other sRNA transcripts produced from intragenic promoters such as the *E. coli* sRNA RbsZ, suggesting that this could be a common mechanism for the production of non-coding transcripts. The expression of RbsZ is driven by an internal promoter to the *rbsB* gene. *rbsB* encodes a periplasmic ribose binding protein required for ribose uptake and impacts the cells ability to carry out chemotaxis towards ribose [67]. RbsZ is a sponge RNA to the sRNA RybB and is discussed further below [68].

### Other mechanisms to produce sponge RNAs

With the exponential increase in the identification of sponge RNAs over the last few years, it is highly conceivable that bacteria have evolved many other mechanisms to produce sponge RNAs. In addition to 3'UTRs [66, 69], it is also conceivable that 5'UTRs could act as sRNAs or sponge RNAs (Figure 1e). In one of the original papers mapping sRNAs in *E. coli*, Vogel *et al* identified several possible sRNAs derived from 5' UTRs of transcripts [70]. Although some of these have since been shown to be riboswitches, it is imaginable that others could have

alternative regulatory roles. One example is provided by the elegant mechanism by which some bacteria integrate ethanolamine and B12 availability. The metabolism of ethanolamine is carried out by the genes encoded by the *eut* operon. The regulation of this operon was studied in *Enterococcus faecalis* [71] and *L. monocytogenes* [72]. The presence of ethanolamine is sensed by the histidine kinase EutW, that in turn phosphorylates and activates the RNA binding protein EutV. EutV binds to RNA hairpin motifs in the *eut* transcript and prevents the termination of transcription when ethanolamine is present. However, the metabolism of ethanolamine by bacteria requires the coenzyme B12. The expression of the genes required for the synthesis of B12 are controlled by a B12 binding riboswitch. The riboswitch is constitutively transcribed and, when B12 is present, B12 binds to the riboswitch causing an alteration of the RNA structure resulting in termination of transcription. However, when B12 is not present the genes for B12 synthesis are transcribed. This transcription results in a longer transcript being produced for the sRNAs Rli55 (*L. monocytogenes*) [72] and EutX (*E. faecalis*) [71]. These sRNAs form a secondary structure downstream of the riboswitch that sequesters the response regulator protein EutV, preventing the transcription of the genes required for ethanolamine metabolism. Therefore, this non-coding RNA is acting as an RNA sponge for the regulatory protein. This elegant mechanism allows the cell to integrate the signals of both ethanolamine and B12 availability.

An alternative system for generating potential regulatory RNA fragments has been described by Dar and Sorek [73]. They described a set of non-coding RNAs they termed decay-generated noncoding RNAs [73] that have been excised from protein coding sequences (Figure 1g). These sequences show an altered degradation rate to their parental mRNA and many have been predicted to interact with the RNA chaperones Hfq and ProQ. These sequences included the sponge RNA SroC. Finally, it is highly conceivable that an sRNA-antisense (as) sRNA pair could act via a sponge RNA mechanism (Figure 1h), whereby the sRNA is expressed under one condition and the as-sRNA under another. Under conditions where the sRNA and as-sRNA are coexpressed, repression or derepression of the sRNA regulon would occur, depending on whether the sRNA acts positively (stabilises) or negatively (blocks) on its targets. This kind of mechanism has already been described for type I toxin-antitoxin systems where by the antitoxin is transcribed antisense to the toxin.

The toxin and antitoxin RNAs interact resulting in translation of the toxin being blocked, the complex being targeted for degradation or in some cases both [74].

## **Mechanisms of action of sponge RNAs**

Table 1 describes many RNA species that interact with sRNAs and affect their ability to act on their mRNA targets. Many terms have been used to describe these RNAs, including sponge RNA [27], ceRNA [14], RNA decoy [26], anti-sRNA [43], RNA predator [24] and RNA trap [75]. This diversity of nomenclature impedes effective literature searches, but it is also intriguing that bacteria have evolved so many different mechanisms to inhibit sRNA activity by titration. The two main mechanisms are depicted in Figure 2 and are discussed below.

### **The sponge RNA target sequence**

Sponge RNAs can interact with their target sRNA in two general ways, either by binding to the same region(s) used by the sRNA to interact with mRNAs or to other regions of the sRNA. However, the result is likely to be similar, namely the targeted sRNA will either be sequestered and held in an inactive form, or the sRNA or sRNA-sponge complex will be degraded (Figure 2a). An unusual example of a sponge RNA that is capable of executing both of these mechanisms is RosA of *B. subtilis*. The interaction of RosA with FsrA results in sequestration, whereas its interaction with RoxS results in degradation of the sRNA [55]. In a  $\Delta$ RosA strain the sRNAs FsrA and RoxS become deregulated. This is reflected in the proteome of the  $\Delta$ RosA strain where targets of FsrA and RoxS were shown to be reduced in their levels [55].

sRNAs can contain one or multiple seed regions, the sections of the RNA that are involved in the interaction with their mRNA targets. A sponge RNA can act to block an sRNA by providing a complementary sequence to the sRNA seed regions. Alternatively, the interaction between the sponge RNA and the sRNA can be at a distance to the seed region. The *Salmonella* sRNA GcvB is one of the best studied sRNAs to date and contains two different seed regions R1 and R3 [39, 61]. GcvB has a very short RNA half-life and an extensive regulon, targeting 1% of all genes in *Salmonella* including those involved in amino acid metabolism [8]. Deletion of *gcvB* results in a faster doubling time than the wild-type



strain when the mutant is grown on media containing tryptone as its nitrogen source, which is abundant in peptides [27]. The deletion of *gcvB* removes the post-transcriptional repression on several peptide permeases that are responsible for uptake of the peptides. As previously stated, one target of GcvB is the 5'UTR of *gltI* from the *gltIJKL* operon encoding the Glutamate/Aspartate ABC transporter, resulting in translational repression of this gene [39]. Interaction between GcvB and *gltI* takes place in seed region R1 [39]. The second GcvB binding site base-pairs to the previously mentioned sponge RNA SroC that is produced from the intergenic spacer between *gltI* and *gltJ* as a result of early termination of transcription and processing by RNase E. Deletion of *sroC* results in a much slower doubling time than the wild-type when grown on the media containing tryptone as the nitrogen source. This is due to the loss of repression of SroC on GcvB in this condition [27] through the interaction of SroC with GcvB at the R3 seed region and a short sequence at the base of stem loop one (SL1) [27, 61] (Table 1). SroC acts as a sponge RNA on a second sRNA MgrR, via a different pairing region from the one involved in the interaction with GcvB [76]. The interaction between SroC and MgrR results in the derepression of *eptB* encoding an LPS modifying enzyme involved in the resistance to the antimicrobial peptide polymyxin B (Table 1).

Interestingly, to date, GcvB has been shown to interact with two different sRNAs, SroC [27] and AgvB [43] (Table 1) and a third, CyaR has been hypothesised [45]. As described above SroC is a multi-target sponge RNA acting on both GcvB [27] and MgrR [76]. The only known target of AgvB is GcvB. AgvB mimics GcvB mRNA targets [43]. Tree et al called AgvB an anti-sRNA and its name denotes *anti*-sRNA for *GcvB*. The genome of Enterohemorrhagic *E. coli* strains 0157:H7 encode two copies of the sequence coding for AgvB. Strains of *E. coli* 0157:H7 are able to colonise the final few centimetres of the bovine gastrointestinal tract, where the majority of the bacteria multiply in the terminal rectal mucus. A mutant deleted for both copies of the AgvB locus was shown to be less competitive in its ability to use mucus as a growth medium, compared to two standard laboratory media [43]. The similarity of the effects of SroC and AgvB, suggest it is due to the amino acid availability in the bovine gastrointestinal tract that leads to the requirement of the post-transcriptional downregulation of GcvB. As such, AgvB contains the consensus target sequence for the R1 seed sequence of GcvB. An example of a mRNA target of GcvB is *dppA* that encodes the dipeptide transporter (Table 1). GcvB and *dppA* interact using the R1 seed region. In the

absence of GcvB, the absence of AgvB was shown to have no effect on *dppA* levels. However, in the presence of GcvB and the absence of AgvB, levels of *dppA* remained low. Mutation of the consensus target sequence in AgvB was also sufficient to allow GcvB repression of *dppA* translation, suggesting that, as for SroC interaction of GcvB with AgvB leads to derepression of the GcvB regulon. The third GcvB interacting partner CyaR was identified through CLASH analysis of Hfq-bound RNAs [45]. CyaR also acts as a classic sRNA regulating mRNA targets [65]. Like AgvB, CyaR is also likely to interact with the R1 seed sequence.

The number of regulatory elements present to control the activity of a single sRNA highlights how important fine-tuning gene expression is for the fitness of the cell. An intriguing question that could be addressed is whether the binding of a sponge RNA to one seed sequence, still allows an sRNA like GcvB to regulate a different subset of targets *via* its second seed sequence if the interaction did not result in degradation of the sRNA. This would add another level to the complexity of post-transcriptional regulation by sponge RNAs.

### **Predatory regulation**

Figuerola-Bossi and Bossi introduced the concept of a sponge RNA being a “predator” [24]. Many sRNAs direct either the target RNA or the sRNA:target complex for degradation by RNases. The very basic concept of a sponge RNA is that it “soaks up” the sRNA with which it interacts. However, if the sponge RNA can, not only efficiently capture its target sRNA, but also promote the destruction of the sRNA, it is not only a sponge RNA, but also a predator. The classic example of a predatory RNA is the intergenic *chbBC* region that titrates the sRNA ChiX away from the *chiP* mRNA. However, this regulation is entirely dependent on the stoichiometry between the ChiX sRNA and the *chbBC* intergenic region [26]. When levels of the *chbBC* mRNA are low compared to ChiX i.e. in the absence of chitosugars, ChiX is able to target this mRNA for degradation. However, when chitosugars are present and the levels of *chbBC* are in excess of the levels of ChiX, then ChiX is targeted for degradation.

The question remains, does every “sponge RNA” that targets its prey for degradation act as a predator or is a true predator a class of its own regulatory RNA? In many cases the

downstream effect on the sRNA has not been determined (Table 1). If a sponge RNA that promotes the degradation of its target is classified as a predator, then under the conditions tested the *B. subtilis* RNA RosA appears to fit the criteria for being both a sponge RNA and a predator [55]. As described previously, one of its targets, FsrA, is sequestered, while the other, RoxS, is degraded. However, RosA appears to be a more multifunctional RNA than has been described for other sponge or predator RNAs, as it is able to alter the level of processing of RoxS to RoxS (D), a shorter form of RoxS, where 20 nucleotides from the 5' end are removed by RNase Y [6]. The absence of RosA also decreased the processing of RoxS to RoxS D, raising the possibility that the RoxS-RosA complex might be able to disassociate to enable RoxS D to regulate its second set of targets [55].

This increased complexity in whether a sponge RNA simply acts to remove a sRNA from the regulatory pool, by blocking its ability to interact with its mRNA target, or promotes its degradation, leads to many questions that have yet to be answered in the field of post-transcriptional gene regulation. If the interaction of the two RNAs results in sequestration, what subsequently happens to this pair of RNAs? Is there possibility for the pair to disassociate, possibly by an RNA helicase if conditions change? Or do interacting pairs of RNAs cycle on and off of each other, with partner swapping depending on relative affinities. Excess levels of the sponge RNA ensures that the sRNA remains blocked, but changes in the stoichiometry of the sRNA, sponge RNA and mRNA target(s) would push the equilibrium in an alternative direction.

### **Level and Threshold of expression**

In regulation of a target by an sRNA the effect is always a question of concentration of either the sRNA or the target. Usually, this change in concentration results from an altered condition that changes the way in which a transcriptional regulator is functioning to either activate or block transcription. As discussed above the concentration of the sRNA ChiX and the *chbBC* transcript is important for whether ChiX regulates *chiP* mRNA or it is sponged out and degraded. It is also a question of RNA stability. Both RoxS and RosA of *B. subtilis* are always expressed, however, which RNA is in excess depends on the conditions [32] (Table 1). RoxS is a very stable sRNA. RosA is regulated by the carbon catabolite control protein CcpA. In conditions where RosA is not repressed by CcpA i.e. the absence of a catabolic

carbon source such as glucose or malate, the half-life of RoxS is approximately 13 minutes, however this increases to over 45 minutes in a  $\Delta$ RosA strain. If the balance is tipped in favour of increased RoxS and decreased RosA transcription, such as upon the addition of malate, then RoxS is able to regulate its target regulon. However, in the absence of malate, transcription will favour increased levels of RosA, which will form a complex with RoxS. Therefore reducing the half-life of RoxS and blocking its interaction with its target regulon, ensuring metabolism efficiently continues down the correct pathway [55].

Control of transcription is not always tight. Regulons of genes controlled by a particular transcription factor are often continually expressed at a low level. In *E. coli*, at least one sponge RNA, the 3'ETS<sup>leuZ</sup>, has evolved to remove the result of this low-level transcriptional noise of the two small RNAs RybB and RyhB (Figure 2b and Table 1) [28], only allowing the regulatory function of these two sRNAs to act when their expression reaches a specific threshold above the expression level of the sponge RNA. tRNA genes are expressed constitutively at a constant level. Processing of the *glyW-cysT-leuZ* releases the 3'ETS<sup>leuZ</sup>. In unstressed cells the sRNAs RybB and RyhB are sequestered by 3'ETS<sup>leuZ</sup> and the levels of their respective targets are unaffected. However, under either iron stress for RybB and envelope stress for RyhB, transcription of these two sRNAs greatly increases. The ratio between the sponge RNA and the sRNA tips in favour of the sRNA and the excess sRNA is free to regulate its downstream targets. As with concentration of an RNA in sRNA-mRNA target regulation, this appears to be a common theme in sponge RNAs. Therefore, controlling the regulatory activity of the sRNA is always a question of threshold between the sponge RNA and its target sRNA, but there are multiple mechanisms by which to achieve this.

## **The global RNA regulatory network**

Bacteria sense their environment at all times and reprogram their gene expression in response to detected changes. A major factor is nutrient availability. Global and local changes ensure that bacteria can make the most of the nutrients that are available and ensure their survival. For example, under starvation conditions, some Firmicutes have the ability to sporulate and remain dormant until conditions improve. A theme that emerges

from the analysis of the known sponge RNAs is the central role they appear to play in balancing metabolism (Table 1). The Gram-negative Enterobacteriaceae and Gram-positive Firmicutes have many overlapping pathways that are similarly regulated at the level of transcription. The same appears to be true for sRNA networks, where several sRNAs have evolved that regulate similar pathways in different bacteria. For example, Fur regulated sRNAs that participate in the iron sparing response to reduce levels of non-essential iron binding proteins are commonly identified [77, 78] and others that are activated by preferred carbon sources, such as Spot 42 in *E. coli* [79] and RoxS in *B. subtilis* [5]. Although the biology of these organisms is quite different, there are key points in their metabolic pathways that clearly need to be controlled, and RNA regulation has been selected as one of the best ways to achieve this goal. Patterns in the evolution of sponge RNAs also appear as more are identified. An intriguing example of sRNA evolution in the Firmicutes is the C-rich region containing RoxS/RsaE sRNA that is conserved in *B. subtilis* and *S. aureus*. However, different CcpA-regulated, G-rich region containing sponge RNAs have evolved to balance the level of these sRNAs. RosA in *B. subtilis* has been shown to reduce levels of RoxS [55], whereas in *S. aureus*, RsaE is known to interact with RsaI, but the mechanism of action is yet to be established [62].

Regulatory RNAs are often studied in isolation, in one condition of interest, where they are deleted or overexpressed and the effect on their targets is monitored. However, global analysis experiments such as CRAC [43], CLASH [45, 46] and RIL-seq [44] have highlighted the fact that RNAs can behave quite differently between conditions. A network of five different sRNAs in *E. coli* act in a condition-specific manner to regulate each other and their targets. ArcZ blocks the activity of CyaR during transition phase [45]. In *Salmonella* ArcZ is also maximally expressed during transition phase, although unlike *E. coli*, expression of ArcZ is detectable during the exponential growth phase [64]. Upon overexpression of the ArcZ sRNA from a multicopy plasmid in stationary phase the levels of over 16% of *Salmonella* transcripts were shown to be altered, many of which have yet to be explained [64]. Some of these changes could feasibly be due to ArcZ acting as a sponge RNA on CyaR and other sRNAs. CyaR and GcvB were shown to interact, likely through their known seed regions, but it has yet to be shown which sRNA is acting as the sponge [45]. As described above GcvB is also sponged by SroC and, in EHEC strains, by the phage encoded sponge RNA AgvB [43].

Therefore, the pattern of repression or derepression of the regulons of specific sRNAs will depend on the combination of different sponge RNAs that are expressed. In addition, like protein regulators, sRNAs have been shown to have a certain hierarchy in the control of their regulons [80, 81]. Bobrovskyy et al., recently studied the SgrS regulon in *E. coli*, involved in the stress response to sugar-phosphate stress. By analysing each of the targets, they identified that the regulatory hierarchy was influenced by features of each sRNA–mRNA pair, the molecular mechanisms of regulation and the role of accessory factors such as Hfq and RNase E [81]. In future research, it will be intriguing to determine how sponge RNAs fit into this hierarchy. Do they always represent the most favoured interaction or can they split the regulon in to a part that is repressed and another part that is derepressed?

In the Enterobacteriaceae a second RNA chaperone, ProQ, was identified through the use of GRAD-seq in *Salmonella* [67] and in studying the development of bacterial competence in *Legionella pneumophila* [82]. ProQ, like the well-studied RNA binding protein Hfq, is able to bind sRNAs. ProQ had previously been proposed to have a role as a matchmaker of sRNA–mRNA pairs [83], however a comparative study of Hfq and ProQ using RIL-seq has shown that this may not be its only function [68]. The study identified a significant proportion of ProQ bound RNA pairs also associated with Hfq. One of these pairs was identified to be the sRNAs RbsZ and RybB, of which RbsZ was shown to be a sponge of the  $\sigma^E$  dependent RybB (Table 1). Hfq is required for the RbsZ mediated degradation of RybB by RNase III. However, interaction of this sponge RNA-sRNA pair with ProQ results in inhibition of this degradation. RbsZ is transcribed from an internal promoter to the *rbsB* gene that is part of the ribose catabolism operon [67]. Under some conditions RbsZ and the end of the *rbs* transcripts are processed to produce a shorter form of RbsZ, denoted RbsZ-S [68]. The full length RbsZ bound at a higher level to ProQ, whereas the processed RbsZ-S bound at a higher level to Hfq. The RIL-seq data also identified chimeras between RybB and the 5' end of the *rbsB* transcript on both Hfq and ProQ. RbsZ-S acts to sponge the sRNA RybB to prevent downregulation of *rbsB*. However, when RybB levels increase above those of RbsZ then RybB is free to downregulate *rbsB*. Melamed et al suggested that this in turn might lead to reduced levels of RbsZ-S and increased active RybB, thus amplifying the negative regulation [68]. It is intriguing what the physiological role of crosstalk between the ribose regulon and

the  $\sigma^E$  dependent sRNA may be, but again highlights the complexity of the post-transcriptional network.

Although less is known about RNA regulation in Gram-positive bacteria, two recent papers in *B. subtilis* [55] and *S. aureus* [62] highlight the fact that RNA networks are also likely in play. RsaI (alternative name RsaOG [84]) is an intriguing sRNA, as identification of its targets by MAPS not only showed that it interacts with multiple mRNAs, but also three different sRNAs, RsaG, RsaD and RsaE (homologue to *B. subtilis* RoxS). However, it interacts with these two groups of RNAs using different seed regions. The mRNAs interact with the unpaired CU rich regions, whereas sRNAs, which all contain C-rich regions are predicted to interact with the G rich regions of RsaI. RsaI is also capable of forming a trimeric complex with RsaG and three different mRNAs *in vitro* [62]. RsaI and RsaG interaction has no effect on the target recognition and regulation of RsaI targets. The RsaG regulon remains to be defined, but it could be predicted that the interaction with RsaI blocks RsaG interacting with its mRNA targets, resulting in a derepression of the RsaG regulon. Therefore, RsaI could be a sponge RNA for RsaG.

## Conclusion and Future Outlook

Sponge RNAs are a diverse group of non-coding RNAs globally present across the prokaryotic and eukaryotic kingdoms. The increase in our ability to carry out both global and targeted RNA interaction studies has widened our view on the types of RNAs that can be classed as sponge RNAs. This is only likely to increase given the type and flexibility now available within the design and analysis of RNA-seq experiments, as was described in this review. These targeted and global studies will enable us to fully understand all the players that are functioning in the post-transcriptional regulatory network. This is likely to include RNAs with multiple functions such as the tRNA intergenic spacer regions that have been described for *E. coli* [28].

There are still further improvements to these methods to be made and in interpreting the data that is generated from them. It is likely that improvements to the molecular biology methods that result in the RNA-seq libraries would enable the identification of more

interactions. For example, in methods such as CLASH, RIL-seq and Modified CLASH, the efficiency of both the cross-linking and ligation steps could be improved. The analysis of the data that results from RNA-seq experiments producing chimeric RNAs is highly complex and there are multiple pipelines that have tried to tackle this challenge. Improvements in mapping the chimeric reads to the genome sequence is key to identifying the interacting RNAs, and key to this is a well annotated genome where the transcript start, stop and processing sites are precisely annotated.

The downstream statistical analysis needs to be carefully conceived to increase the confidence in identifying real RNA interactions. Many of these techniques by their nature create problems with background. However due to the nature of RNA interactions and RNA turnover often being highly efficient, many chimeric reads in these data sets are likely to be rare. Therefore, the statistics used in identifying both false positives and false negatives needs to be carefully considered. The additional use of sequence interaction software such as IntaRNA [11, 85] has been shown to be beneficial in identifying both false positives and false negatives [46, 55]. When a list of RNA interactions is produced from the analysis of the RNA-seq data, at present it is challenging without further validation to determine how the identified targets of a particular sRNA are likely to act. For example, is the interaction between an sRNA and a possible target acting to block or enhance translation? Or is it, indeed acting as a *bone fide* sponge RNA?

If we were able to fully describe all of the bona fide interactions within a particular bacterium, the resulting studies to understand the regulatory logic mediated by all the sRNAs and sponge RNAs in the network may not be possible. Notably, we would likely need to understand the dynamics of the regulation of the RNAs themselves, and parameters such as their half-life and relative affinities to understand the putative hierarchies of the interactions. We can then hope to address questions as to whether the circuitry is over-elaborate and reflective of the tinkering hand of evolution [86] or more closely reflect optimal control structures with energetic and response times optimized in a manner that maximizes fitness. Consistent with tinkering (bricolage), it is intriguing that intervening spacer regions have been recruited as RNA sponges.



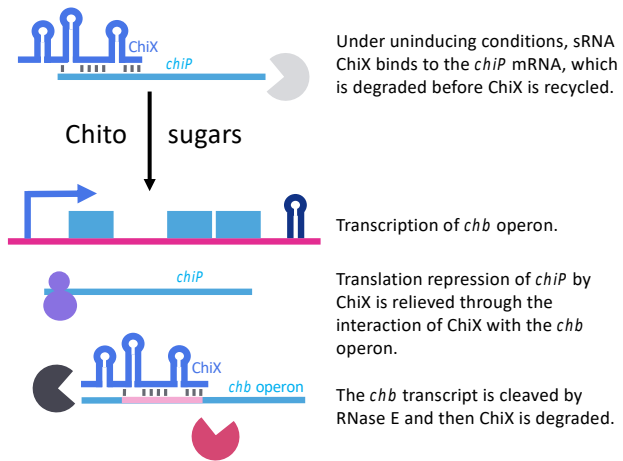
This review has discussed examples of how sponge RNAs act not only as stand-alone regulators of sRNAs, but can be multipurpose, blocking the function of sRNAs in addition to interacting with their own mRNA targets. This adds an increased level of complexity to the regulatory network where crosstalk between sRNAs can affect the levels of the different sRNA regulons either positively or negatively. This fine tuning of gene expression is a mechanism in how bacteria gain fitness benefits and out-compete their neighbours. The mechanisms these molecules utilise have already been highlighted for use in eukaryotic systems for not only understanding how networks function [87], but also in molecular medicine, where in one trial the hepatitis C virus was inhibited in cell culture using a sponge RNA [88]. Sponge RNAs offer similar possibilities in prokaryotic systems not only in synthetic biology, but also in increasing our development of novel RNA based therapies.

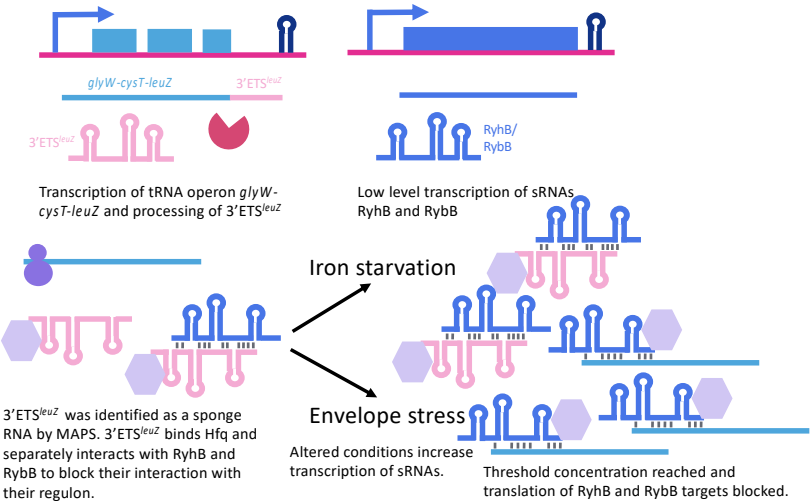
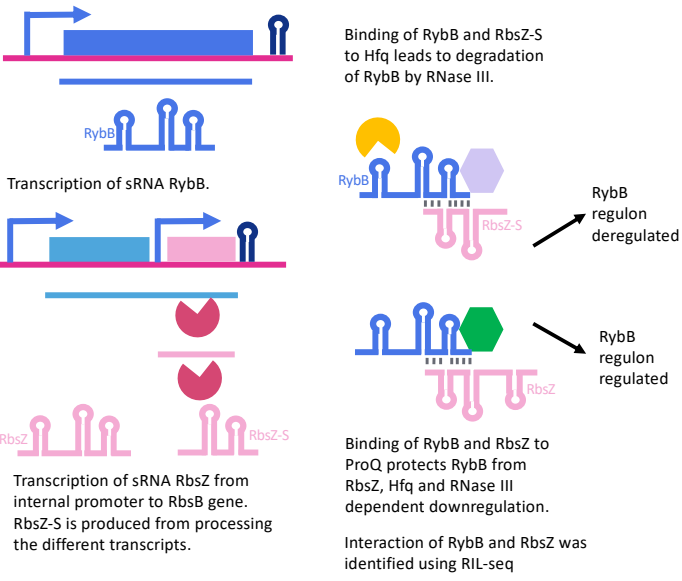
## **Acknowledgements**

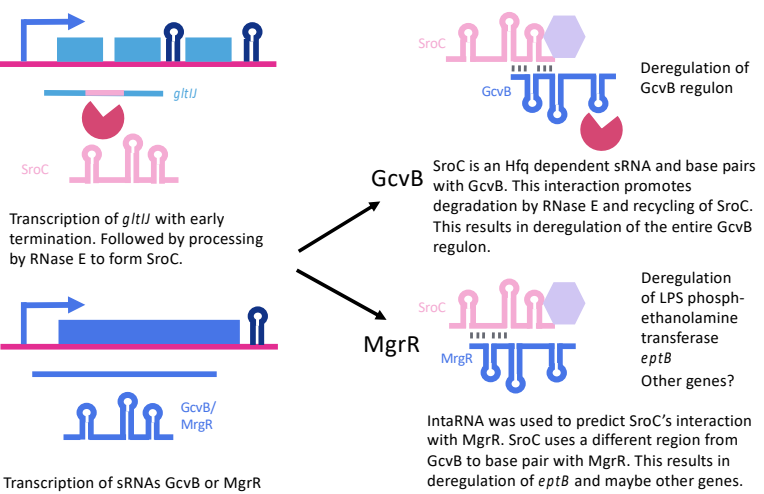
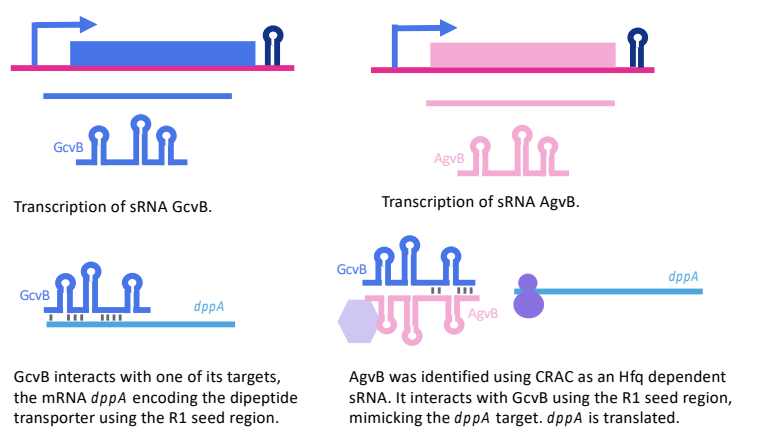
Thank you to Sylvain Durand, Ciarán Condon, Laurence Hurst and Thomas Krogh Kallehauge for useful discussions and critical reading of the manuscript. Thank you to the two reviewers for their helpful comments. This work was supported by the University of Bath.

**Table 1 The bacterial Sponge RNAs and their regulatory effects on their target(s)**

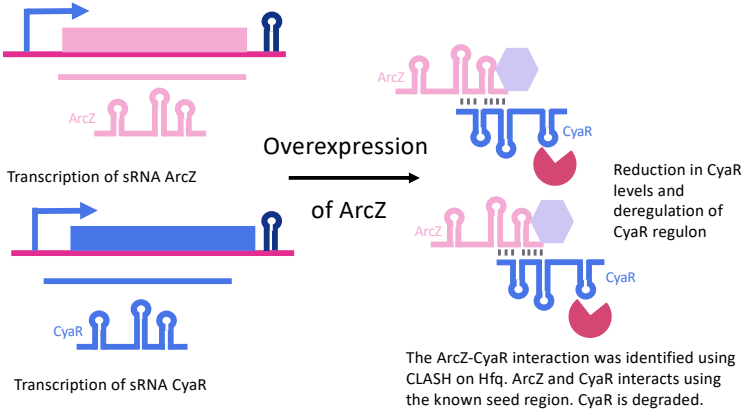
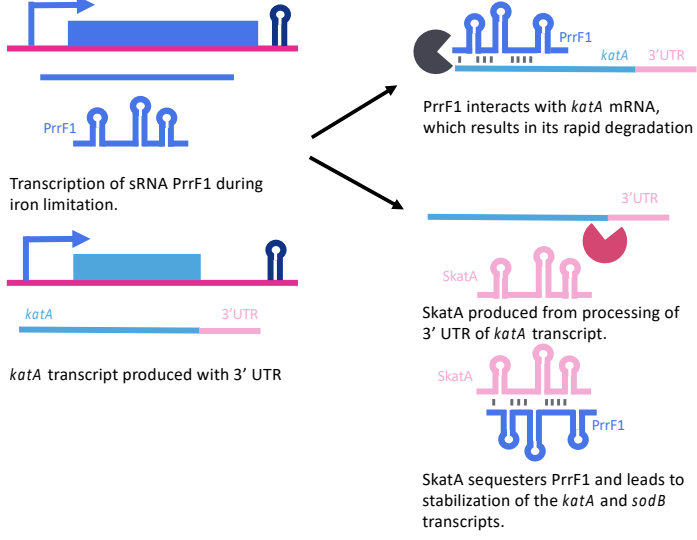
The coding strand of the DNA is depicted as dark pink, promoter sequences as blue arrows, genes as turquoise blocks, mRNAs as turquoise lines, terminator sequences dark blue, sponge RNA coding regions pink blocks and sponge RNAs as pink lines. Ribosomes are shown in purple, Hfq in lilac, ProQ in green. RNases are depicted in red (RNase E), orange (RNase III), black and grey (exonuclease).

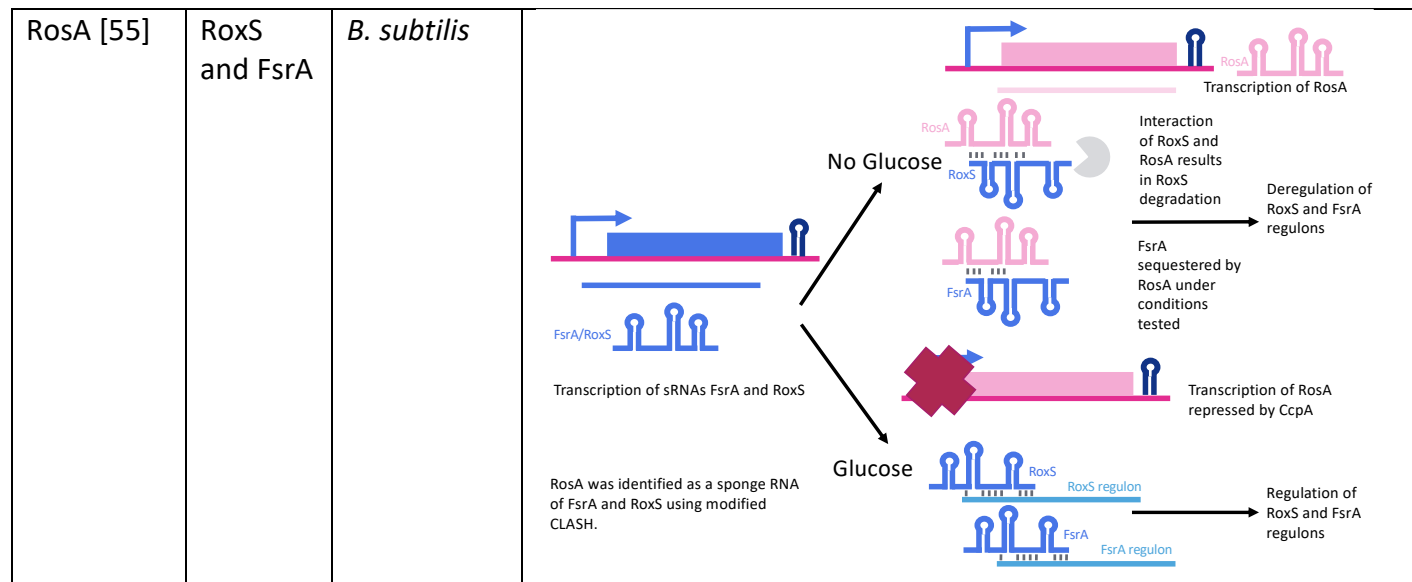
Sponge RNA	sRNA target(s)	Species	Mechanism of action
Intergenic region of <i>chbBC</i> transcript [22, 26, 89]	ChiX	Salmonella and <i>E. coli</i>	 <p>Under uninducing conditions, sRNA ChiX binds to the <i>chiP</i> mRNA, which is degraded before ChiX is recycled.</p> <p>Chito sugars</p> <p>Transcription of <i>chb</i> operon.</p> <p>Translation repression of <i>chiP</i> by ChiX is relieved through the interaction of ChiX with the <i>chb</i> operon.</p> <p>The <i>chb</i> transcript is cleaved by RNase E and then ChiX is degraded.</p>

<p>3'ETS<sup>leuZ</sup> [28]</p>	<p>RyhB and RybB</p>	<p>Enterobacte riaceae</p>	 <p>Transcription of tRNA operon <i>glyW-cysT-leuZ</i> and processing of 3'ETS<sup>leuZ</sup></p> <p>Low level transcription of sRNAs RyhB and RybB</p> <p>Iron starvation</p> <p>Envelope stress</p> <p>Altered conditions increase transcription of sRNAs.</p> <p>Threshold concentration reached and translation of RyhB and RybB targets blocked.</p> <p>3'ETS<sup>leuZ</sup> was identified as a sponge RNA by MAPS. 3'ETS<sup>leuZ</sup> binds Hfq and separately interacts with RyhB and RybB to block their interaction with their regulon.</p>
<p>RbsZ and RbsZ (S) [68]</p>	<p>RybB</p>	<p><i>E. coli</i></p>	 <p>Binding of RybB and RbsZ-S to Hfq leads to degradation of RybB by RNase III.</p> <p>Transcription of sRNA RybB.</p> <p>RybB regulon deregulated</p> <p>Transcription of sRNA RbsZ from internal promoter to RbsB gene. RbsZ-S is produced from processing the different transcripts.</p> <p>RybB regulon regulated</p> <p>Binding of RybB and RbsZ to ProQ protects RybB from RbsZ, Hfq and RNase III dependent downregulation.</p> <p>Interaction of RybB and RbsZ was identified using RIL-seq</p>

SroC [27]	GcvB [8] and MgrR [76]	Enterobacteriaceae	 <p>Transcription of <i>gltI</i> with early termination. Followed by processing by RNase E to form SroC.</p> <p>SroC is an Hfq dependent sRNA and base pairs with GcvB. This interaction promotes degradation by RNase E and recycling of SroC. This results in deregulation of the entire GcvB regulon.</p> <p>Deregulation of LPS phosphoethanolamine transferase <i>eptB</i> Other genes?</p> <p>IntaRNA was used to predict SroC's interaction with MgrR. SroC uses a different region from GcvB to base pair with MgrR. This results in deregulation of <i>eptB</i> and maybe other genes.</p> <p>Transcription of sRNAs GcvB or MgrR</p>
AgvB [43]	GcvB	<i>E. coli</i> EHEC Two copies on SP10 and SP17	 <p>Transcription of sRNA GcvB.</p> <p>GcvB interacts with one of its targets, the mRNA <i>dppA</i> encoding the dipeptide transporter using the R1 seed region.</p> <p>Transcription of sRNA AgvB.</p> <p>AgvB was identified using CRAC as an Hfq dependent sRNA. It interacts with GcvB using the R1 seed region, mimicking the <i>dppA</i> target. <i>dppA</i> is translated.</p>

AsxR [43]	FnrS	<i>E. coli</i> EHEC Stx2Φ prophage	<p>Transcription of sRNA FnrS.</p> <p>Transcription of sRNA AsxR.</p> <p>FnrS interacts with one of its targets, the mRNA <i>chuS</i> encoding the heme oxidase using seed region.</p> <p>AsxR was identified using CRAC as an Hfq dependent sRNA. It interacts with FnrS using its terminator stem loop. FnrS is degraded and <i>chuS</i> is translated.</p>
PspH [44]	Spot 42	<i>E. coli</i>	<p>Transcription of PspH from its own promoter</p> <p>Overexpression of PspH</p> <p>Deregulation of Spot 42 regulon</p> <p>PspH was identified to interact with Spot 42 using RIL-seq. The known Spot 42 seed region participates in the interaction.</p>

ArcZ [45]	CyaR	<i>E. coli</i>	 <p>Transcription of sRNA ArcZ</p> <p>Transcription of sRNA CyaR</p> <p>Overexpression of ArcZ</p> <p>Reduction in CyaR levels and deregulation of CyaR regulon</p> <p>The ArcZ-CyaR interaction was identified using CLASH on Hfq. ArcZ and CyaR interacts using the known seed region. CyaR is degraded.</p>
SkatA [58]	PrrF1	<i>P. aeruginosa</i>	 <p>Transcription of sRNA PrrF1 during iron limitation.</p> <p><i>katA</i> transcript produced with 3' UTR</p> <p>PrrF1 interacts with <i>katA</i> mRNA, which results in its rapid degradation</p> <p>SkatA produced from processing of 3' UTR of <i>katA</i> transcript.</p> <p>SkatA sequesters PrrF1 and leads to stabilization of the <i>katA</i> and <i>sodB</i> transcripts.</p>



**Table 2 Methods to identify RNA-RNA interactions including sponge RNAs**

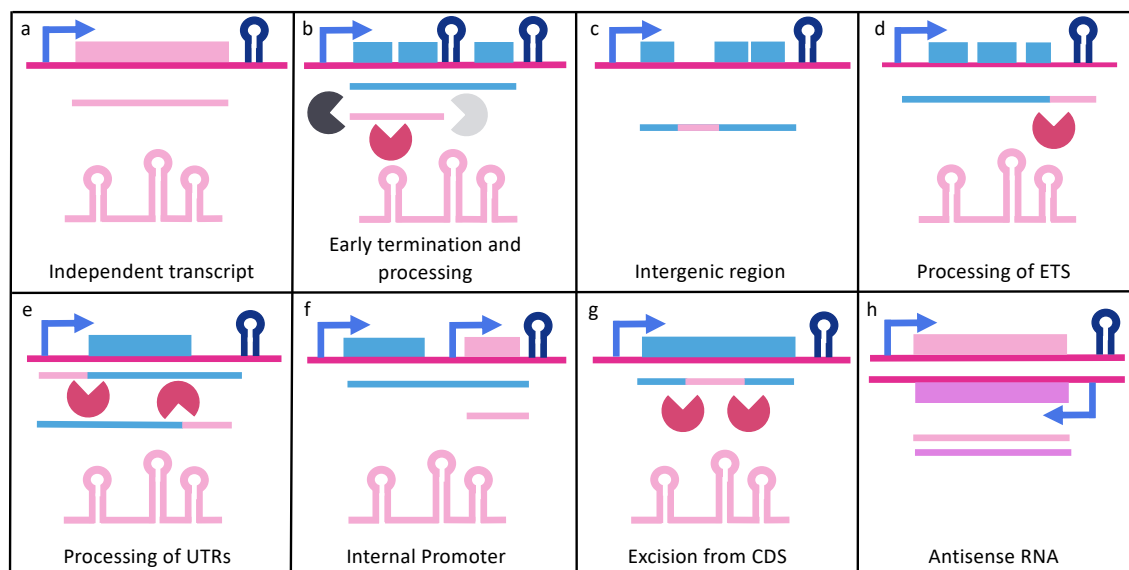
Technique Name	Methodology	Advantages and Disadvantages	Key References
<b>Protein dependent methods</b>			
RIP-seq (RNA immunoprecipitation followed by RNA-seq)	An antibody to the protein of interest can be used or the protein is tagged with a suitable purification tag such as His, Flag or Spa. At the required point in growth bacteria are harvested, disrupted to release the cell contents before being incubated with the protein or tag specific antibody. The antibody bound protein is purified using protein A Sepharose beads before the RNA is extracted.	- This is the simplest of the protocols that allow identification of RNAs associated with RBPs. – A list of RNAs associated with the RBP is generated and further computational and molecular work is required to determine which RNAs in the list interact.	[41, 66]
CLIP-seq (Cross-linking Immunoprecipitation)  CLIP-seq has many variations [42]	This is a modification to the RIP-seq protocol where UV-C (254 nm) is used to cross-link RNAs to the RBP of interest, before the IP. The IP is carried out under stringent conditions with washing steps including high-salt buffers and ionic detergents that ensures only the cross-linked protein-RNA complex is purified. To extract the RNA from the RBP, proteinase K (PK) is used to cleave the protein cross-linked to the RNA under denaturing conditions. Releasing the RNA into solution with a peptide that remains on the RNA at the site of crosslinking.	- Increased stringency of the protein purification reduces background. Identification of the point of RNA-protein cross-linking. - No information as to which RNAs interact with each other, so therefore like RIP-seq computational and molecular work is required to determine which RNAs in the list interact.	[90, 91]
CRAC (Crosslinking and Analysis of cDNA)	This is a variation of CLIP-seq where a two-step affinity purification under denaturing conditions. This ensures other RNA binding proteins and free RNAs that are not cross-linked to the protein of interest are removed. In the RNA extraction step urea and SDS are used to enhance the activity of PK.	- Increased purification and stringency ensures the sample is not contaminated by other proteins or RNAs. - No information as to which RNAs interact with each other, so therefore like RIPseq computational and molecular work is required to determine	[43, 92, 93]



		which RNAs in the list interact.	
CLASH (Cross-linking, ligation and sequencing of hybrids)	This is a variation of CLIP and CRAC where not only is UV cross-linking used to pinpoint the exact point where the protein and RNA interact, but the interacting RNAs are ligated by proximity ligation to create chimeric RNAs. The purification is carried out under stringent conditions. The RNA-seq data analysis pipeline identifies the two RNAs present within the chimera. Approaches such as determining the folding energy of the chimera helps determine true interactions.	<ul style="list-style-type: none"> <li>- Purification carried out under stringent conditions. Ligation of RNAs allows interacting pairs of RNAs to be identified.</li> <li>- Some interactions may be missed due to ligation favouring some RNA pairs over others. Ligation may occur between non-regulator pairs due to the proximity of the RNAs on the protein.</li> </ul>	[45, 46, 94, 95]
RIL-seq (RNA interaction by ligation and sequencing)	RIL-seq is a similar method to CLASH, where UV cross-linking is used and RNAs are ligated to form chimeras on the RBP being studied, but the purification is carried out under native conditions.	<ul style="list-style-type: none"> <li>- Unlike many RNA-RBP protocols there is no requirement for radiolabelling.</li> <li>- Some interactions may be missed due to ligation favouring some RNA pairs over others. Ligation may occur between non-regulator pairs due to the proximity of the RNAs on the protein.</li> </ul>	[44, 68, 96]
<b>Protein independent methods</b>			
Modified CLASH	This method is similar to RBP based CLASH, but instead of the RBP the psoralen AMT is used to cross-link interacting RNAs together. The bacteria are grown as required before the addition of AMT that is taken up by the cell and crosslinking with UV at 365 nm. The RNA is harvested, rRNA depleted, single stranded overhangs are removed with RNase digestion and the resulting ends of the interacting RNAs are ligated to form chimeric RNAs. Non-ligated RNAs are removed using further RNase treatment	<ul style="list-style-type: none"> <li>- No requirement of genetic modification. Enables identification of RNA interactions without knowledge of an RBP.</li> <li>- Efficiency of crosslinking and ligation of interacting RNAs is low. The psoralen AMT only allows crosslinks of interactions between uridine residues. False positives in the data set are likely to be high</li> </ul>	[51, 52, 54, 55, 97]

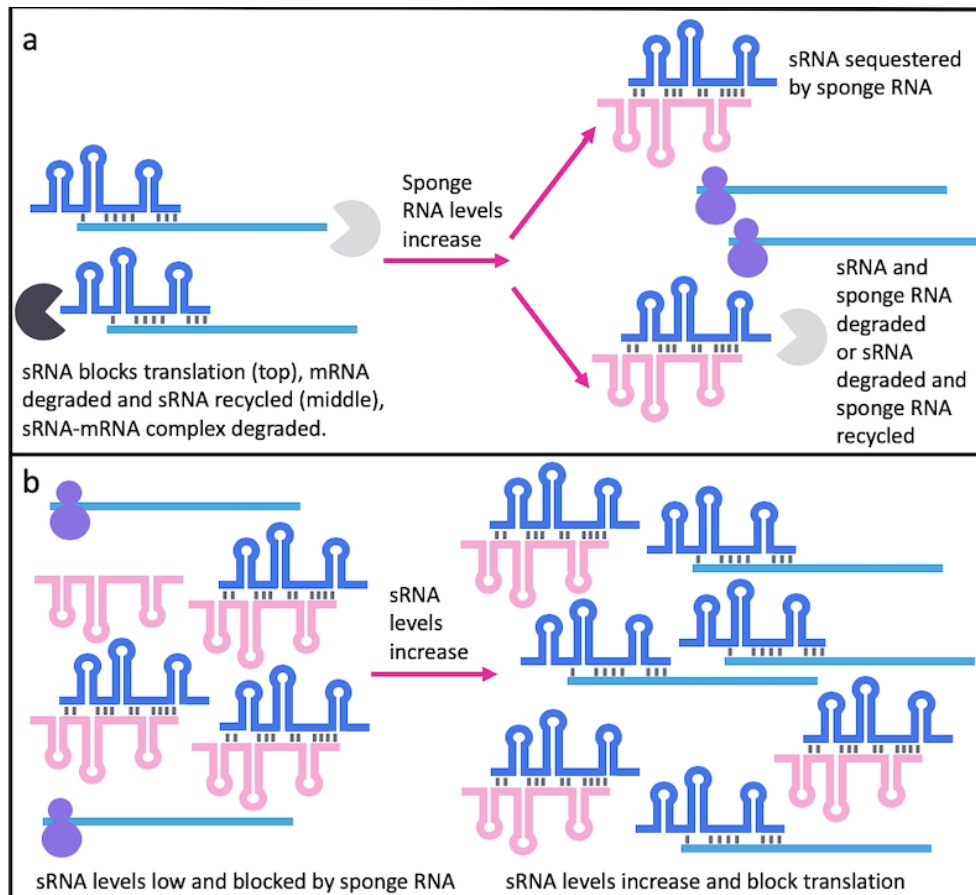
	and the crosslinking reversed using UV 254 nm. RNAs are sequenced and analysed to identify chimeric RNAs.	due to the ligation step being in free solution rather than on a protein as for RBP-CLASH.	
GRIL-seq (Global sRNA Target Identification by Ligation and Sequencing)	The sRNA to be studied is overexpressed from an inducible promoter, followed by overexpression of the T4 RNA-ligase. The sRNA of interest is purified from the pool of RNA using complementary poly(A) tailed oligonucleotides. This allows capture of the sRNA including the chimeras on oligo dT magnetic beads. Following RNA-seq the identity of the RNA in the chimeras are determined.	<ul style="list-style-type: none"> <li>- High proportion of reads map to sRNA being studied. Requires over expression of both the sRNA and the RNA ligase <i>in vivo</i>.</li> <li>- Focuses on one sRNA at a time. Ligation requires a 5' monophosphate and therefore will only allow identification of processed RNAs where the 5' triphosphate has been removed.</li> </ul>	[58]
Hi-GRIL-seq	Very similar to GRIL-seq, but allows the identification of RNA-RNA interactions at a global level rather than focussing on one. The T4-RNA ligase is overexpressed to ligate interacting RNAs before the RNA is harvested, rRNA depleted and RNA-seq carried out to allow the interacting RNAs to be identified.	<ul style="list-style-type: none"> <li>- Global RNA interaction study.</li> <li>- Requires over expression of the RNA ligase <i>in vivo</i>. The efficiency of the ligation is very low. Likely high false positive identification rate due to ligation occurring <i>in vivo</i> and no processing of RNA ends to ensure ligation of ends of RNA are in close proximity.</li> </ul>	[59]
MAPS (MS2 affinity purification coupled with RNA-sequencing)	The MS2 stem loop aptamer is added to either the 5' or 3' end of the RNA of interest and is expressed <i>in vivo</i> . The tagged RNA is then harvested in the condition of study before being purified by affinity chromatography. The resulting RNA is sequenced and compared to a non-tagged control. Targets of the sRNA will be enriched.	<ul style="list-style-type: none"> <li>- High number of reads for interactions with sRNA being studied.</li> <li>- Requires over-expression of the sRNA. Addition of the MS2 tag to the sRNA may result in altered expression or secondary structure of the sRNA and therefore not allowing identification of</li> </ul>	[28, 60, 98]

		targets. One sRNA can be studied at a time.	
--	--	---	--



**Figure 1 – Mechanisms to produce sponge RNAs**

The coding strand of the DNA is depicted as dark pink, promoter sequences as blue arrows, genes as turquoise blocks, mRNAs as turquoise lines, terminator sequences dark blue, sponge RNA coding regions pink blocks and sponge RNA as pink lines. **a)** Independent transcript, produced from its own promoter sequence and is likely to have a specific terminator sequence. Examples of independent transcripts acting as sponge RNAs *RosA B. subtilis* and *AsxR* expressed from the *Stx2Φ* prophage **b)** Early termination and processing of transcript as depicted by the coloured RNases. This processing can be 5' (dark grey), 3' (light grey) or within the RNA sequence (burgundy). Some transcripts have multiple terminator sequences. Under certain conditions the first terminator may form and transcription is terminated. The transcript is then processed and can act as a sponge RNA. *SroC* is processed by RNase E from the *gltII* transcript and targets *GcvB*. **c)** Intergenic region. The intergenic region of the *chbBC* transcript mimics the sRNA binding site of the sRNA *ChiX* that under inducing conditions targets *ChiX* for degradation. **d)** Processing of External Transcribed Spacer (ETS). Similar to processing the 3' UTR by an endoribonuclease. The 3' ETS<sup>leuZ</sup> of *E. coli* is an example. **e)** Processing of UTR. Many genes have long regions transcribed before (5' UTR) or after the coding sequence (3' UTR). These regions can be cleaved an endoribonuclease to release a functional regulatory RNA. The 3' UTR of the *katA* gene of *P. aeruginosa* is processed to release the sponge RNA *SkatA*. A sponge RNA from a 5' UTR has yet to be identified. **f)** Internal Promoter. This has been described for the 3' UTR of *PspG* which results in the sponge RNA *PspH*. However, the promoter could be located anywhere in the gene. The following have yet to be identified as acting as sponge RNAs, but are likely examples for future investigation. **g)** Excision from protein coding sequence. **h)** Antisense RNA.



**Figure 2 – Mechanisms by which sponge RNAs can regulate sRNAs**

mRNAs are depicted as turquoise lines, ribosomes are in purple, sRNAs as blue lines, sponge RNA as pink lines, RNases black and grey.

**a)** Sponge RNA is expressed in a condition specific manner. The sponge RNA can either mimic the mRNA sequence that the sRNA targets. When conditions change, the levels of the sponge RNA increase and it interacts with the sRNA to block its interaction with the mRNA target. Upon binding with the sRNA, the complex is either targeted for degradation by RNases or the sponge RNA sequesters the sRNA and the complex is stably maintained in the cell. **b)** Threshold concentration. The sponge RNA is present at a constant level and interacts with the sRNA, which is expressed at a low level. When conditions change, the rate of transcription of the sRNA increases and pushes its concentration above the level of the sponge RNA. The free sRNA is able to interact with its mRNA targets and block the translation and or target the complex for degradation by RNases.

## References

- [1] T. Mizuno, M.Y. Chou, M. Inouye, A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA), *Proceedings of the National Academy of Sciences of the United States of America*, 81 (1984) 1966-1970.
- [2] N. Delihas, Discovery and characterization of the first non-coding RNA that regulates gene expression, micF RNA: A historical perspective, *World journal of biological chemistry*, 6 (2015) 272-280.
- [3] E. Holmqvist, C. Unoson, J. Reimegard, E.G. Wagner, A mixed double negative feedback loop between the sRNA MicF and the global regulator Lrp, *Molecular microbiology*, 84 (2012) 414-427.
- [4] C.P. Corcoran, D. Podkaminski, K. Papenfort, J.H. Urban, J.C. Hinton, J. Vogel, Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA, *Molecular microbiology*, 84 (2012) 428-445.
- [5] S. Durand, F. Braun, A.C. Helfer, P. Romby, C. Condon, sRNA-mediated activation of gene expression by inhibition of 5'-3' exonucleolytic mRNA degradation, *eLife*, 6 (2017).
- [6] S. Durand, F. Braun, E. Lioliou, C. Romilly, A.C. Helfer, L. Kuhn, N. Quittot, P. Nicolas, P. Romby, C. Condon, A nitric oxide regulated small RNA controls expression of genes involved in redox homeostasis in *Bacillus subtilis*, *PLoS genetics*, 11 (2015) e1004957.
- [7] R.A. Mars, P. Nicolas, M. Ciccolini, E. Reilman, A. Reder, M. Schaffer, U. Mader, U. Volker, J.M. van Dijl, E.L. Denham, Small regulatory RNA-induced growth rate heterogeneity of *Bacillus subtilis*, *PLoS genetics*, 11 (2015) e1005046.
- [8] C.M. Sharma, K. Papenfort, S.R. Pernitzsch, H.J. Mollenkopf, J.C. Hinton, J. Vogel, Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA, *Molecular microbiology*, 81 (2011) 1144-1165.
- [9] K.S. Frohlich, S. Gottesman, Small Regulatory RNAs in the Enterobacterial Response to Envelope Damage and Oxidative Stress, *Microbiology spectrum*, 6 (2018).
- [10] P.R. Wright, A.S. Richter, K. Papenfort, M. Mann, J. Vogel, W.R. Hess, R. Backofen, J. Georg, Comparative genomics boosts target prediction for bacterial small RNAs, *Proceedings of the National Academy of Sciences of the United States of America*, 110 (2013) E3487-3496.
- [11] P.R. Wright, J. Georg, M. Mann, D.A. Sorescu, A.S. Richter, S. Lott, R. Kleinkauf, W.R. Hess, R. Backofen, CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains, *Nucleic acids research*, 42 (2014) W119-123.
- [12] B. Tjaden, TargetRNA: a tool for predicting targets of small RNA action in bacteria, *Nucleic acids research*, 36 (2008) W109-113.
- [13] B. Tjaden, Computational identification of sRNA targets, *Methods in molecular biology* (Clifton, N.J.), 905 (2012) 227-234.
- [14] L. Salmena, L. Poliseno, Y. Tay, L. Kats, P.P. Pandolfi, A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language?, *Cell*, 146 (2011) 353-358.
- [15] M.S. Ebert, J.R. Neilson, P.A. Sharp, MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells, *Nature methods*, 4 (2007) 721-726.
- [16] J.M. Franco-Zorrilla, A. Valli, M. Todesco, I. Mateos, M.I. Puga, I. Rubio-Somoza, A. Leyva, D. Weigel, J.A. Garcia, J. Paz-Ares, Target mimicry provides a new mechanism for regulation of microRNA activity, *Nature genetics*, 39 (2007) 1033-1037.
- [17] R.O. Bak, J.G. Mikkelsen, miRNA sponges: soaking up miRNAs for regulation of gene expression, *Wiley interdisciplinary reviews. RNA*, 5 (2014) 317-333.

- [18] D.W. Thomson, M.E. Dinger, Endogenous microRNA sponges: evidence and controversy, *Nature reviews. Genetics*, 17 (2016) 272-283.
- [19] L. Bossi, N. Figueroa-Bossi, Competing endogenous RNAs: a target-centric view of small RNA regulation in bacteria, *Nature reviews. Microbiology*, 14 (2016) 775-784.
- [20] K. Papenfort, C.K. Vanderpool, Target activation by regulatory RNAs in bacteria, *FEMS microbiology reviews*, 39 (2015) 362-378.
- [21] T. Romeo, P. Babitzke, Global Regulation by CsrA and Its RNA Antagonists, *Microbiology spectrum*, 6 (2018).
- [22] N. Figueroa-Bossi, M. Valentini, L. Malleret, F. Fiorini, L. Bossi, Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target, *Genes & development*, 23 (2009) 2004-2015.
- [23] M.S. Azam, C.K. Vanderpool, Talk among yourselves: RNA sponges mediate cross talk between functionally related messenger RNAs, *The EMBO journal*, 34 (2015) 1436-1438.
- [24] N. Figueroa-Bossi, L. Bossi, Sponges and Predators in the Small RNA World, *Microbiology spectrum*, 6 (2018).
- [25] M.P. Grull, E. Masse, Mimicry, deception and competition: The life of competing endogenous RNAs, *Wiley interdisciplinary reviews. RNA*, 10 (2019) e1525.
- [26] J. Plumbridge, L. Bossi, J. Oberto, J.T. Wade, N. Figueroa-Bossi, Interplay of transcriptional and small RNA-dependent control mechanisms regulates chitosugar uptake in *Escherichia coli* and *Salmonella*, *Molecular microbiology*, 92 (2014) 648-658.
- [27] M. Miyakoshi, Y. Chao, J. Vogel, Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA, *The EMBO journal*, 34 (2015) 1478-1492.
- [28] D. Lalaouna, M.C. Carrier, S. Semsey, J.S. Brouard, J. Wang, J.T. Wade, E. Masse, A 3' external transcribed spacer in a tRNA transcript acts as a sponge for small RNAs to prevent transcriptional noise, *Molecular cell*, 58 (2015) 393-405.
- [29] L. Argaman, R. Hershberg, J. Vogel, G. Bejerano, E.G.H. Wagner, H. Margalit, S. Altuvia, Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*, *Current Biology*, 11 (2001) 941-950.
- [30] K.M. Wassarman, F. Repoila, C. Rosenow, G. Storz, S. Gottesman, Identification of novel small RNAs using comparative genomics and microarrays, *Genes & development*, 15 (2001) 1637-1651.
- [31] E. Rivas, R.J. Klein, T.A. Jones, S.R. Eddy, Computational identification of noncoding RNAs in *E. coli* by comparative genomics, *Current Biology*, 11 (2001) 1369-1373.
- [32] P. Nicolas, U. Mader, E. Dervyn, T. Rochat, A. Leduc, N. Pigeonneau, E. Bidnenko, E. Marchadier, M. Hoebeke, S. Aymerich, D. Becher, P. Bisicchia, E. Botella, O. Delumeau, G. Doherty, E.L. Denham, M.J. Fogg, V. Fromion, A. Goelzer, A. Hansen, E. Hartig, C.R. Harwood, G. Homuth, H. Jarmer, M. Jules, E. Klipp, L. Le Chat, F. Lecointe, P. Lewis, W. Liebermeister, A. March, R.A. Mars, P. Nannapaneni, D. Noone, S. Pohl, B. Rinn, F. Rugheimer, P.K. Sappa, F. Samson, M. Schaffer, B. Schwikowski, L. Steil, J. Stulke, T. Wiegert, K.M. Devine, A.J. Wilkinson, J.M. van Dijl, M. Hecker, U. Volker, P. Bessieres, P. Noirot, Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*, *Science*, 335 (2012) 1103-1106.
- [33] A. Toledo-Arana, O. Dussurget, G. Nikitas, N. Sesto, H. Guet-Revillet, D. Balestrino, E. Loh, J. Gripenland, T. Tiensuu, K. Vaitkevicius, M. Barthelemy, M. Vergassola, M.A. Nahori, G. Soubigou, B. Regnault, J.Y. Coppee, M. Lecuit, J. Johansson, P. Cossart, The *Listeria* transcriptional landscape from saprophytism to virulence, *Nature*, 459 (2009) 950-956.

- [34] U. Mader, P. Nicolas, M. Depke, J. Pane-Farre, M. Debarbouille, M.M. van der Kooi-Pol, C. Guerin, S. Derozier, A. Hiron, H. Jarmer, A. Leduc, S. Michalik, E. Reilman, M. Schaffer, F. Schmidt, P. Bessieres, P. Noirot, M. Hecker, T. Msadek, U. Volker, J.M. van Dijk, *Staphylococcus aureus* Transcriptome Architecture: From Laboratory to Infection-Mimicking Conditions, *PLoS genetics*, 12 (2016) e1005962.
- [35] U. Nagalakshmi, Z. Wang, K. Waern, C. Shou, D. Raha, M. Gerstein, M. Snyder, The transcriptional landscape of the yeast genome defined by RNA sequencing, *Science*, 320 (2008) 1344-1349.
- [36] C.M. Sharma, S. Hoffmann, F. Darfeuille, J. Reignier, S. Findeiss, A. Sittka, S. Chabas, K. Reiche, J. Hackermuller, R. Reinhardt, P.F. Stadler, J. Vogel, The primary transcriptome of the major human pathogen *Helicobacter pylori*, *Nature*, 464 (2010) 250-255.
- [37] D. Dar, M. Shamir, J.R. Mellin, M. Koutero, N. Stern-Ginossar, P. Cossart, R. Sorek, Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria, *Science*, 352 (2016) aad9822.
- [38] J.B. Lalanne, J.C. Taggart, M.S. Guo, L. Herzel, A. Schieler, G.W. Li, Evolutionary Convergence of Pathway-Specific Enzyme Expression Stoichiometry, *Cell*, 173 (2018) 749-761.e738.
- [39] C.M. Sharma, F. Darfeuille, T.H. Plantinga, J. Vogel, A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites, *Genes & development*, 21 (2007) 2804-2817.
- [40] J. Vogel, B.F. Luisi, Hfq and its constellation of RNA, *Nature reviews. Microbiology*, 9 (2011) 578-589.
- [41] A. Sittka, S. Lucchini, K. Papenfort, C.M. Sharma, K. Rolle, T.T. Binnewies, J.C. Hinton, J. Vogel, Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq, *PLoS genetics*, 4 (2008) e1000163.
- [42] F.C.Y. Lee, J. Ule, Advances in CLIP Technologies for Studies of Protein-RNA Interactions, *Molecular cell*, 69 (2018) 354-369.
- [43] J.J. Tree, S. Granneman, S.P. McAteer, D. Tollervey, D.L. Gally, Identification of bacteriophage-encoded anti-sRNAs in pathogenic *Escherichia coli*, *Molecular cell*, 55 (2014) 199-213.
- [44] S. Melamed, A. Peer, R. Faigenbaum-Romm, Y.E. Gatt, N. Reiss, A. Bar, Y. Altuvia, L. Argaman, H. Margalit, Global Mapping of Small RNA-Target Interactions in Bacteria, *Molecular cell*, 63 (2016) 884-897.
- [45] I.A. Iosub, M. Marchiorretto, B. Sy, S. McKellar, K.J. Nieken, R.W. van Nues, J.J. Tree, G. Viero, S. Granneman, Hfq CLASH uncovers sRNA-target interaction networks involved in adaptation to nutrient availability, *bioRxiv*, (2019) 481986.
- [46] S.A. Waters, S.P. McAteer, G. Kudla, I. Pang, N.P. Deshpande, T.G. Amos, K.W. Leong, M.R. Wilkins, R. Strugnell, D.L. Gally, D. Tollervey, J.J. Tree, Small RNA interactome of pathogenic *E. coli* revealed through crosslinking of RNase E, *The EMBO journal*, (2016).
- [47] M. Dambach, I. Irnov, W.C. Winkler, Association of RNAs with *Bacillus subtilis* Hfq, *PLoS one*, 8 (2013) e55156.
- [48] H. Hammerle, F. Amman, B. Vecerek, J. Stulke, I. Hofacker, U. Blasi, Impact of Hfq on the *Bacillus subtilis* transcriptome, *PLoS one*, 9 (2014) e98661.
- [49] T. Rochat, O. Delumeau, N. Figueroa-Bossi, P. Noirot, L. Bossi, E. Dervyn, P. Boulloc, Tracking the Elusive Function of *Bacillus subtilis* Hfq, *PLoS one*, 10 (2015) e0124977.
- [50] C. Bohn, C. Rigoulay, P. Boulloc, No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*, *BMC microbiology*, 7 (2007) 10.



- [51] J.G. Aw, Y. Shen, A. Wilm, M. Sun, X.N. Lim, K.L. Boon, S. Tapsin, Y.S. Chan, C.P. Tan, A.Y. Sim, T. Zhang, T.T. Susanto, Z. Fu, N. Nagarajan, Y. Wan, In Vivo Mapping of Eukaryotic RNA Interactomes Reveals Principles of Higher-Order Organization and Regulation, *Molecular cell*, 62 (2016) 603-617.
- [52] Z. Lu, Q.C. Zhang, B. Lee, R.A. Flynn, M.A. Smith, J.T. Robinson, C. Davidovich, A.R. Gooding, K.J. Goodrich, J.S. Mattick, J.P. Mesirov, T.R. Cech, H.Y. Chang, RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure, *Cell*, 165 (2016) 1267-1279.
- [53] E. Sharma, T. Sterne-Weiler, D. O'Hanlon, B.J. Blencowe, Global Mapping of Human RNA-RNA Interactions, *Molecular cell*, 62 (2016) 618-626.
- [54] T. Liu, K. Zhang, S. Xu, Z. Wang, H. Fu, B. Tian, X. Zheng, W. Li, Detecting RNA-RNA interactions in *E. coli* using a modified CLASH method, *BMC genomics*, 18 (2017) 343.
- [55] S. Durand, A. Callan-Sidat, J. McKeown, S. Li, G. Kostova, J.R.H. Feraud, M.T. Alam, A. Millard, C. Constantinidou, C. Condon, E.L. Denham, Novel regulation from novel interactions: Identification of an RNA sponge that controls the levels, processing and efficacy of the RoxS riboregulator of central metabolism in *Bacillus subtilis*, *bioRxiv*, (2019) 814905.
- [56] A. Gaballa, H. Antelmann, C. Aguilar, S.K. Khakh, K.B. Song, G.T. Smaldone, J.D. Helmann, The *Bacillus subtilis* iron-sparing response is mediated by a Fur-regulated small RNA and three small, basic proteins, *Proceedings of the National Academy of Sciences of the United States of America*, 105 (2008) 11927-11932.
- [57] G.T. Smaldone, O. Revelles, A. Gaballa, U. Sauer, H. Antelmann, J.D. Helmann, A global investigation of the *Bacillus subtilis* iron-sparing response identifies major changes in metabolism, *Journal of bacteriology*, 194 (2012) 2594-2605.
- [58] K. Han, B. Tjaden, S. Lory, GRIL-seq provides a method for identifying direct targets of bacterial small regulatory RNA by in vivo proximity ligation, *Nature microbiology*, 2 (2016) 16239.
- [59] Y.F. Zhang, K. Han, C.E. Chandler, B. Tjaden, R.K. Ernst, S. Lory, Probing the sRNA regulatory landscape of *P. aeruginosa*: post-transcriptional control of determinants of pathogenicity and antibiotic susceptibility, *Molecular microbiology*, 106 (2017) 919-937.
- [60] M.C. Carrier, D. Lalaouna, E. Masse, A game of tag: MAPS catches up on RNA interactomes, *RNA biology*, 13 (2016) 473-476.
- [61] D. Lalaouna, A. Eyraud, A. Devinck, K. Prevost, E. Masse, GcvB small RNA uses two distinct seed regions to regulate an extensive targetome, *Molecular microbiology*, 111 (2019) 473-486.
- [62] D. Bronesky, E. Desgranges, A. Corvaglia, P. Francois, C.J. Caballero, L. Prado, A. Toledo-Arana, I. Lasa, K. Moreau, F. Vandenesch, S. Marzi, P. Romby, I. Caldelari, A multifaceted small RNA modulates gene expression upon glucose limitation in *Staphylococcus aureus*, *The EMBO journal*, (2019).
- [63] A. Tomasini, K. Moreau, J. Chicher, T. Geissmann, F. Vandenesch, P. Romby, S. Marzi, I. Caldelari, The RNA targetome of *Staphylococcus aureus* non-coding RNA RsaA: impact on cell surface properties and defense mechanisms, *Nucleic acids research*, 45 (2017) 6746-6760.
- [64] K. Papenfort, N. Said, T. Welsink, S. Lucchini, J.C.D. Hinton, J. Vogel, Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA, *Molecular microbiology*, 74 (2009) 139-158.

- [65] N. De Lay, S. Gottesman, The Crp-activated small noncoding regulatory RNA CyaR (RyeE) links nutritional status to group behavior, *Journal of bacteriology*, 191 (2009) 461-476.
- [66] Y. Chao, K. Papenfort, R. Reinhardt, C.M. Sharma, J. Vogel, An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs, *The EMBO journal*, 31 (2012) 4005-4019.
- [67] M.K. Thomason, T. Bischler, S.K. Eisenbart, K.U. Forstner, A. Zhang, A. Herbig, K. Nieselt, C.M. Sharma, G. Storz, Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*, *Journal of bacteriology*, 197 (2015) 18-28.
- [68] S. Melamed, P.P. Adams, A. Zhang, H. Zhang, G. Storz, RNA-RNA Interactomes of ProQ and Hfq Reveal Overlapping and Competing Roles, *Molecular cell*, (2019).
- [69] Y. Chao, L. Li, D. Girodat, K.U. Forstner, N. Said, C. Corcoran, M. Smiga, K. Papenfort, R. Reinhardt, H.J. Wieden, B.F. Luisi, J. Vogel, In Vivo Cleavage Map Illuminates the Central Role of RNase E in Coding and Non-coding RNA Pathways, *Molecular cell*, 65 (2017) 39-51.
- [70] J. Vogel, V. Bartels, T.H. Tang, G. Churakov, J.G. Slagter-Jager, A. Huttenhofer, E.G. Wagner, RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria, *Nucleic acids research*, 31 (2003) 6435-6443.
- [71] S. DebRoy, M. Gebbie, A. Ramesh, J.R. Goodson, M.R. Cruz, A. van Hoof, W.C. Winkler, D.A. Garsin, Riboswitches. A riboswitch-containing sRNA controls gene expression by sequestration of a response regulator, *Science*, 345 (2014) 937-940.
- [72] J.R. Mellin, M. Koutero, D. Dar, M.A. Nahori, R. Sorek, P. Cossart, Riboswitches. Sequestration of a two-component response regulator by a riboswitch-regulated noncoding RNA, *Science*, 345 (2014) 940-943.
- [73] D. Dar, R. Sorek, Bacterial Noncoding RNAs Excised from within Protein-Coding Transcripts, *mBio*, 9 (2018).
- [74] S. Masachis, F. Darfeuille, Type I Toxin-Antitoxin Systems: Regulating Toxin Expression via Shine-Dalgarno Sequence Sequestration and Small RNA Binding, *Microbiology spectrum*, 6 (2018).
- [75] J. Vogel, An RNA trap helps bacteria get the most out of chitosugars, *Molecular Microbiology*, 73 (2009) 737-741.
- [76] L.G. Acuna, M.J. Barros, D. Penaloza, P.I. Rodas, D. Paredes-Sabja, J.A. Fuentes, F. Gil, I.L. Calderon, A feed-forward loop between SroC and MgrR small RNAs modulates the expression of eptB and the susceptibility to polymyxin B in *Salmonella Typhimurium*, *Microbiology*, 162 (2016) 1996-2004.
- [77] S. Chareyre, P. Mandin, Bacterial Iron Homeostasis Regulation by sRNAs, *Microbiology spectrum*, 6 (2018).
- [78] A.G. Oglesby-Sherrouse, E.R. Murphy, Iron-responsive bacterial small RNAs: variations on a theme, *Metallomics : integrated biometal science*, 5 (2013) 276-286.
- [79] S. Durica-Mitic, Y. Gopel, B. Gorke, Carbohydrate Utilization in Bacteria: Making the Most Out of Sugars with the Help of Small Regulatory RNAs, *Microbiology spectrum*, 6 (2018).
- [80] J. Fei, D. Singh, Q. Zhang, S. Park, D. Balasubramanian, I. Golding, C.K. Vanderpool, T. Ha, RNA biochemistry. Determination of in vivo target search kinetics of regulatory noncoding RNA, *Science*, 347 (2015) 1371-1374.

- [81] M. Bobrovskyy, M.S. Azam, J.K. Frandsen, J. Zhang, A. Poddar, X. Ma, T.M. Henkin, T. Ha, C.K. Vanderpool, Determinants of target prioritization and regulatory hierarchy for the bacterial small RNA SgrS, *Molecular microbiology*, (2019).
- [82] L. Attaiech, A. Boughammoura, C. Brochier-Armanet, O. Allatif, F. Peillard-Fiorente, R.A. Edwards, A.R. Omar, A.M. MacMillan, M. Glover, X. Charpentier, Silencing of natural transformation by an RNA chaperone and a multitarget small RNA, *Proceedings of the National Academy of Sciences of the United States of America*, 113 (2016) 8813-8818.
- [83] A. Smirnov, C. Wang, L.L. Drewry, J. Vogel, Molecular mechanism of mRNA repression in trans by a ProQ-dependent small RNA, *The EMBO journal*, 36 (2017) 1029-1045.
- [84] A. Marchais, C. Bohn, P. Boulloc, D. Gautheret, RsaOG, a new Staphylococcal family of highly transcribed non-coding RNA, *RNA biology*, 7 (2010) 116-119.
- [85] A. Busch, A.S. Richter, R. Backofen, IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions, *Bioinformatics (Oxford, England)*, 24 (2008) 2849-2856.
- [86] F. Jacob, Evolution and tinkering, *Science*, 196 (1977) 1161-1166.
- [87] Y. Yuan, B. Liu, P. Xie, M.Q. Zhang, Y. Li, Z. Xie, X. Wang, Model-guided quantitative analysis of microRNA-mediated regulation on competing endogenous RNAs using a synthetic gene circuit, *Proceedings of the National Academy of Sciences of the United States of America*, 112 (2015) 3158-3163.
- [88] I. Jost, L.A. Shalamova, G.K. Gerresheim, M. Niepmann, A. Bindereif, O. Rossbach, Functional sequestration of microRNA-122 from Hepatitis C Virus by circular RNA sponges, *RNA biology*, 15 (2018) 1032-1039.
- [89] M. Overgaard, J. Johansen, J. Moller-Jensen, P. Valentin-Hansen, Switching off small RNA regulation with trap-mRNA, *Molecular microbiology*, 73 (2009) 790-800.
- [90] J. Ule, K. Jensen, A. Mele, R.B. Darnell, CLIP: a method for identifying protein-RNA interaction sites in living cells, *Methods (San Diego, Calif.)*, 37 (2005) 376-386.
- [91] E. Holmqvist, P.R. Wright, L. Li, T. Bischler, L. Barquist, R. Reinhardt, R. Backofen, J. Vogel, Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo, *The EMBO journal*, 35 (2016) 991-1011.
- [92] B. Sy, J. Wong, S. Granneman, D. Tollervey, D. Gally, J.J. Tree, High-Resolution, High-Throughput Analysis of Hfq-Binding Sites Using UV Crosslinking and Analysis of cDNA (CRAC), *Methods in molecular biology (Clifton, N.J.)*, 1737 (2018) 251-272.
- [93] M.T. Bohnsack, D. Tollervey, S. Granneman, Identification of RNA helicase target sites by UV cross-linking and analysis of cDNA, *Methods in enzymology*, 511 (2012) 275-288.
- [94] A.J. Travis, J. Moody, A. Helwak, D. Tollervey, G. Kudla, Hyb: a bioinformatics pipeline for the analysis of CLASH (crosslinking, ligation and sequencing of hybrids) data, *Methods (San Diego, Calif.)*, 65 (2014) 263-273.
- [95] A. Helwak, D. Tollervey, Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH), *Nature protocols*, 9 (2014) 711-728.
- [96] S. Melamed, R. Faigenbaum-Romm, A. Peer, N. Reiss, O. Shechter, A. Bar, Y. Altuvia, L. Argaman, H. Margalit, Mapping the small RNA interactome in bacteria using RIL-seq, *Nature protocols*, 13 (2018) 1-33.
- [97] C.A. Weidmann, A.M. Mustoe, K.M. Weeks, Direct Duplex Detection: An Emerging Tool in the RNA Structure Analysis Toolbox, *Trends in biochemical sciences*, 41 (2016) 734-736.
- [98] M.C. Carrier, G. Laliberte, E. Masse, Identification of New Bacterial Small RNA Targets Using MS2 Affinity Purification Coupled to RNA Sequencing, *Methods in molecular biology (Clifton, N.J.)*, 1737 (2018) 77-88.

